LANGUAGE IN INDIA

Strength for Today and Bright Hope for Tomorrow

Volume 3 : 1 January 2003

Editor: M. S. Thirumalai, Ph.D.
Associate Editors: B. Mallikarjun, Ph.D.
Sam Mohanlal, Ph.D.
B. A. Sharada, Ph.D.

PRE-REQUISITES FOR THE PREPARATION OF AN ELECTRONIC THESAURUS FOR A TEXT PROCESSOR IN INDIAN LANGUAGES

S. Rajendran, Ph.D.

---

1. The Role Of Electronic Thesaurus In A Text Processor

The Text-processors/Word-processors, which are mostly meant for documenting text materials, comes handy for the purpose of working papers, books, etc. and editing and printing them to suit ones requirement. The Text processor helps one to write a text into the computer and get the output in the desired format. The text processor is provided with a sophisticated mechanism of writing the text into the computer and manipulates it according to the need of the user. There are a number of Text Processors available in the computer market and a few commonly used processors are WORDSTAR, MS-WORD and WORDPERFECT.

MS-WORD is chosen here for our study. Apart from the facilities provided by edit menu for copying, cutting and pasting, deleting, etc., the processor has a tool menu which provides the following unique facilities:

1. Facilities for spell checking
2. Facility for grammar checking
3. Facility for selecting appropriate and needed vocabulary

Here we are interested in the third facility listed above. This facility is available in the name of thesaurus. Thesaurus tool is available in most of the existing text processors that are meant for English language. The text processors available for Indian languages generally do not have the above-mentioned three facilities as it involves a lot of homework by linguists and language technologists.

The thesaurus tool available in a text processor helps a user in selecting an appropriate equivalent for the lexical item, which has been written in the text, and replacing the already written one by the selected one. When the thesaurus is called for it gives the user with a list of synonyms, antonyms and related words. The user is at his liberty to choose the one of his choices or rejects the suggestion. The MS-word, which is a widely used text processor, gives at least the following thesaurus information:

1. List of replaceable synonymous lexical items.
2. Meaning of the word.
3. Related word or words

For example, if a user invokes the **Thesaurus** for the word he has typed in his text, say, for example, ***electronic***, the dialogue box of the Thesaurus will show the following:

Replaceable Synonyms
automatic, automated, mechanical, computerized, programmed, streamlined, cybernetic

Meaning
automatic (adj.)

Related item
electron

And if one proceeds to look up for the thesaurus information for the word electron the three expected information will be given as follows:

Replaceable synonyms
particle in an atom, atom, negative particle, negation, electrically charged element. neutron, proton, particles

Meaning
particle of an atom (noun)

For the word preparation the following thesaurus information is available:

Replaceable synonyms
arrangement, plan, groundwork, proceeding, readiness, preparedness, adaptation, treatment, incubation

Meaning
arrangement (noun), education (noun), compound (noun), making ready (noun)

Related words
prepare

In this article, we do not propose to evaluate the thesaurus facility available in this text processor for English. We plan to look forward to preparing an Electronic Thesaurus for Text Processing (shortly ETTP) for Indian languages, which, in fact, is more ambitious and complex than the one we have seen above. This will reflect the mental make up, or the psychological make up of the mental lexicon, so that the user can utilize the said thesaurus in whatever way he likes to make use of. Tamil language is taken for this case study. The text processor is so ambitious that suppose one wants to write about a novel centering around a hospital, he will be provided with the lexical items that are related to the hospital situation. This will be a great boon especially in the Indian context, since most writers have difficulty in finding the right word for such conepts in the Indian language they use.

2. Net Work Of Lexical Relations

Understanding the network of lexical relations existing between words (i.e. lexical items) is a prerequisite for a thesaurus in a text processor. The important lexical relations that have to be studied are the following:

1. Synonymy
2. Hyponymy
3. Compatibility
4. Incompatibility

5. Meronymy
6. Morphological relations.

These lexical relations help a thesaurus maker to organize and group lexical items into semantic domains, sub domains and lexical sets. Certain groups of items can be arranged hierarchically, while certain groups deserve different type of arrangement.

2. 1. Synonymy, Hyponymy, Compatibility And Incompatibility

A word acquires its referential meaning in being a member of a semantic domain by the common features it shares with other members in that domain, and having contrasting features, which separate it from other members of the domain. It is the semantic relations among words, such as synonymy, hyponymy, compatibility and incompatibility, which help one to classify and organize words in terms of semantic features or components in a hierarchical or orderly fashion. Componential analysis of meaning is a welcome thing to achieve this mission of relating and classifying lexical items by semantic features.

2.1. Synonymy

The lexical items which have the same meaning or which share same componential features are synonyms and the relationship existing between them is synonymy. Synonymy does not necessarily mean that the items concerned should be identical in meaning, i.e. interchangeable in all contexts. Synonymy can be said to occur if lexical items are close enough in their meaning to allow a choice to be made between them in some contexts, without there being any difference in the meaning of the sentence as a whole. Take, for examples, the words nduul 'book' and puttakam 'book'. avan nduul paTikkiRaan 'he is reading a book' can entail avan puttakam paTikkiRaan 'he is reading a book'. The relation existing between nduul and puttakam is synonymy and nduul and puttakam are synonyms.

2.3. Hyponymy

Hyponymy is the relationship that exists between specific and general lexical items, such that the former is included in the latter. The set of terms, which are hyponyms of same superordinate term, are co-hyponyms. Take for example the lexical items pacu 'cow', erumai 'buffalo' and vilangku 'animal'. itu oru pacu 'this is a cow'and itu oru erumai 'this is a buffalo' unilaterally entail itu oru vilanku 'this is an animal'. The relationship existing between pacu and vilangku and erumai and vilangku is hyponymy and pacu and erumai are co-hyponyms.

2.3. Compatibility

The lexical items that overlap in terms of meaning and do not show systematic include-included relation and have some semantic traits in common, but differ in respect of traits that do not clash are said to be compatible. Take for example the words ***naay*** 'dog' and ***cellappiraaNi*** 'pet'. A dog could be a pet, but neither all pets are not dogs nor all dogs are pets. The relationship existing between ***ndaay*** and ***cellappiraaNi*** is compatible.

2.4. Incompatibility

Incompatibility refers to sets of items where the choice of one item excludes the use of all the other items from that set. Take for example the words ***puunai*** 'cat' and ***ndaay*** 'dog'. ***itu oru puunai*** 'it is a cat' can entail ***itu oru ndaay illai*** 'it is not a dog'. The relation existing between puunai and naay is incompatible. Both come under the superordinate term ***vilanku*** 'animal'. Thus the incompatible items can be co-hyponyms of a superordinate item, that is items which are in incompatible can be related to one another by hyponymous relation. All kinds of oppositions can be included under incompatibility. If the opposition is between two lexical items, it is called binary opposition and if the opposition is between many lexical items it is called many-member opposition.

2.4.1. Binary Opposition

Antonymy, which is often considered as opposite to synonymy, relies on the lexical relation of incompatibility. Lyons (1977:270-290) reserves the term antonymy for a narrow sense, and describes several different kinds of lexical 'opposites'. For him antonyms are

always gradable opposites (e.g. ***peritu*** 'big' : ***ciRitu*** 'small', which are distinguished from complementaries, which are not gradable (e.g. ***aaN*** 'male' : ***peN*** 'female'). Cutting across this distinction of antonyms and complementaries is that drawn between privative and equipollent opposites: a privative opposition is a contrastive relation between two lexical items, one of which denotes some positive property and the other of which the absence of that property (e.g. ***uyartiNai*** 'rational' : ***ahRiNai*** 'irrational') and equipollent opposition is a relation in which each of the contrasting lexical items denotes a positive property
(eg. ***tiTapporuL*** 'solid' : ***tiravapporuL*** 'liquid').

Antonymy
E.g. peritu 'big' : ciRitu 'small'

Complementarity
aaN 'male' : peN 'female'

Privative opposition
ahRiNai 'irratiopnal' : uyartiNai 'rational'

Equipollent opposition
tiTapporuL 'solid' : tiravapporuL 'liquid'

A further opposition to be distinguished from antonymy and complementarity is converseness, exemplified by such pairs as kaNavan 'husband': manaivi 'wife'. Converses typically include the following relations.

1. **Reciprocal Social roles**
maruttuyar 'doctor' : ndooyaaLi 'patient'

2. **Kinship Relations**
appaa 'father' : makan 'son' ammaa 'mother' : makaL 'daughter'

3. **Temporal Relations**
munnar 'before' : pinnar 'after'

4. **Spatial Relations**
meelee 'above' : kiizhee 'below'

5. **Complex Relations**
vaangku 'buy' : vil 'sell'

In addition, there can be defined three further oppositions. They are:

1. **Dircetional Opposition**
vantuceer 'arrive' : puRappaTu 'depart' vaa 'come' : poo 'go'

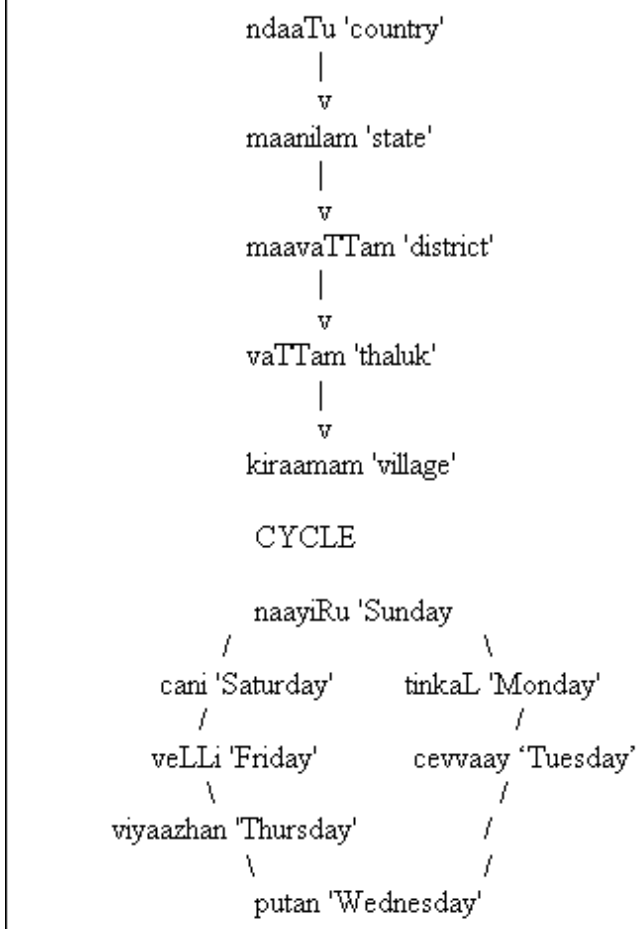
2. **Orthogonal Opposition or Perpendicular Opposition**
vaTakku 'north' : kizhakku 'east' and meeRku 'west' kizhakku 'east' : teRku 'south' and vaTakku 'north'

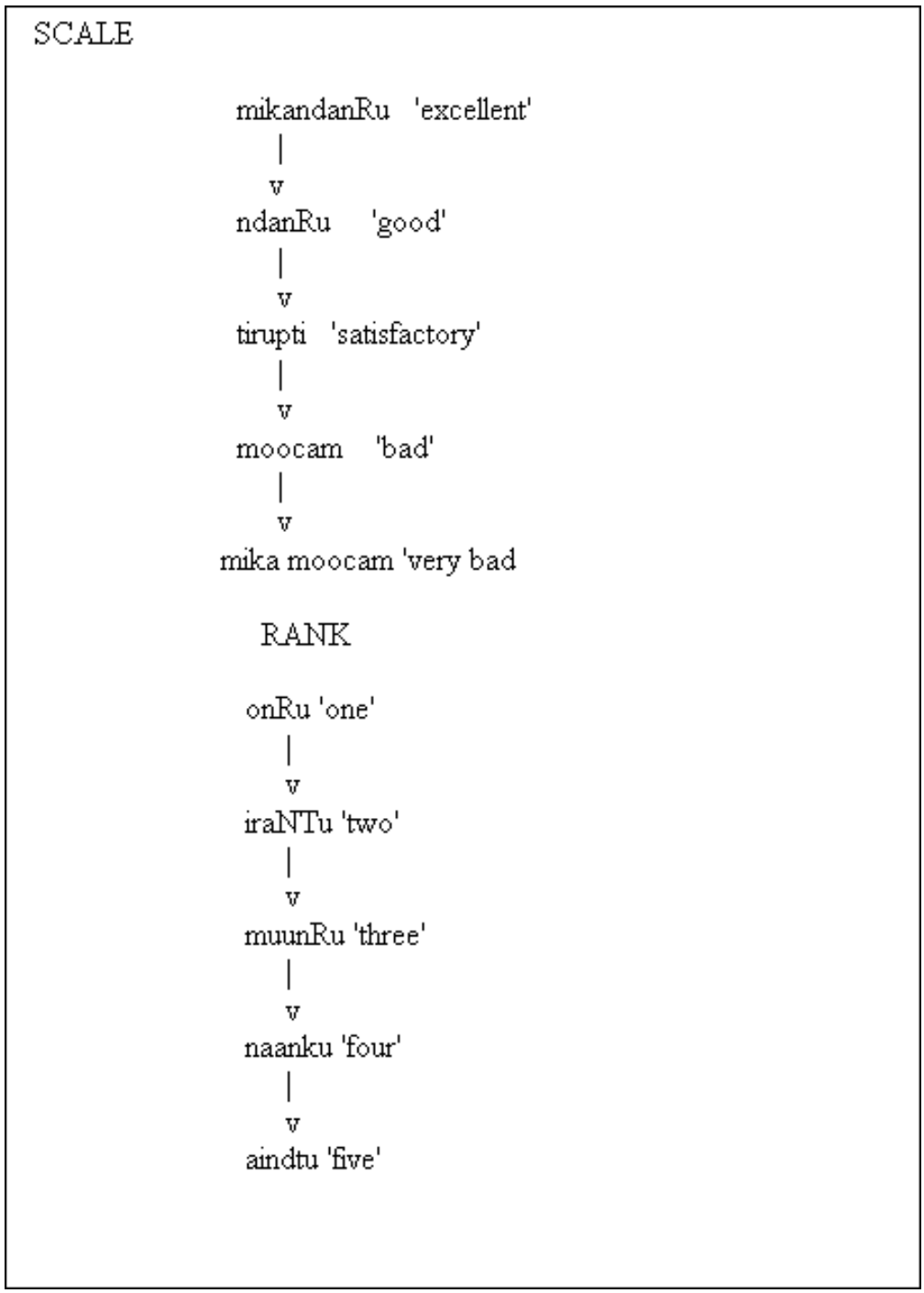
3. **Antipodal Opposition or Diametrical Opposition**
vaTakku 'north' : teRku 'south' kizhakku 'east : meeRku 'west'

2.4.2. Multi-Member Opposition

There are different types of multi-member sets in a language whose lexical relations can be described as incompatible denoting non-binary contrasts as opposed to binary contrasts. Various kinds of ordering are found in multi-member sets of incompatibles, and such sets may be serially or cyclically ordered. In a serially ordered set there are two outer most member and all other lexical items in the set are ordered between them. In a cyclically ordered sets every lexical item is ordered between two others. Among serially ordered sets, scales may be distinguished from ranks according to whether the constituent lexical items are gradable or not. The ordering in scales in terms of incompatibility is typically less strict than it in ranks. The examples given below will stand to explain the difference between them.

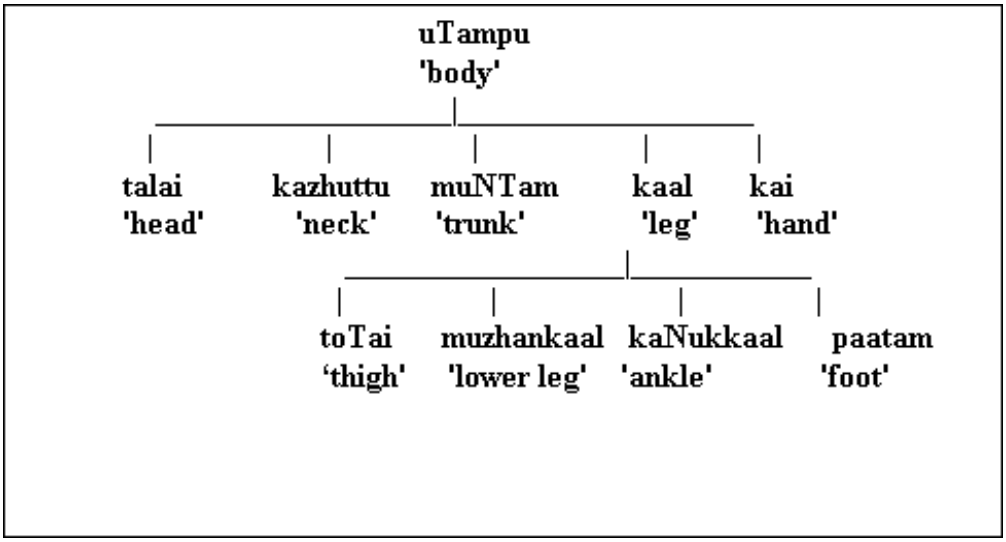
SERIAL


ndaaTu 'country'
maanilam 'state'
maavaTTam 'district'
vaTTam 'thaluk'
kiraamam 'village'
CYCLE
naayiRu 'Sunday
cani 'Saturday'
tinkaL 'Monday'
veLLi 'Friday'
cevvaay 'Tuesday'
viyaazhan 'Thursday'
putan 'Wednesday'



2.5. Meronymy

Meronymy (part-whole relation) also plays an important role in the hierarchical arrangement of lexical items. The division of the human body into parts can serve as a prototype for all part-whole hierarchies:



The superordinate term in the part-whole relationship is holonym and the parts which are subordinate to the superordinate term are meronyms. A holonym may have many meronyms. In the above example uTampu 'body' is a holonym and talai, kazhuttu, muNTam, kaal, kai are meronyms. In the same spirit, kaal is the holonym of meronyms such as toTai 'thigh', 'muzhankaal', and paatam 'foot'.

2.5.1. Morphological Relations

The morphological relations crops up when one attempts to prepare ETTP. This is because the processor has to help a user to replace an inflected word by a synonymous word which is also of the same inflected type. That is, the word kuuRinaan 'said_he' has be replaced by the synonymous word connan 'said_he' and not by col . That means the relation between col 'say' , connaan 'said-he' , colkiRaan 'says_he' and colvaan has to be captured. This requires a morphological analyser and a word generator the preparation of which is not elaborated further in this paper. Also the ETTP has to be designed to give related word or words for the word in the text, the derivational relation has also to be captured. For example, the relation between preparation and prepare has to be captured.

3. Nida's Hierarchical Classification

Nida (1975a) who was concerned with the preparation of a thesaurus dictionary for Greek gives the following as the tentative hierarchical classification of the lexical items (Nida:178-186).

I. Entities
A. Inanimate
1. Natural
a. Geographical
b. Natural substances
c. Flora and plant products
2. Manufactured or constructed entities
a. Artifacts (non constructions)
b. Processed substances: foods, medicines, and perfumes
c. Constructions
B. Animate entities
1. Animals, birds, insects
2. Human beings
3. Supernatural power or beings

II. Events
A. Physical, B. Physiological, C. Sensory, D. Emotive, E. Intellection, G. Communication, G. Association, H. Control, I. Movement, J. Impact, K. Transfer, L. Complex activities, involving a series of movements or actions

III. Abstracts
A. Time, B. Distance, C. Volume, D. Velocity, E. Temperature, F. Color, G. Number, H. Status, I. Religious character, J. Attractiveness, K. Age, L. Truth-falsehood, M. Good-bad, N. Capacity, O. State of health, etc.

IV. Relationals
A. Spatial, B. Temporal, C. Deictic , D. Logical, etc.

This classification is based on referential meanings and it is not possible to obtain one to one correspondence between the semantic domain of classes and the grammatical classes. There are, of course certain parallel between them, "since on some level of the deep structure entries tend to be represented by nouns, events by verbs, abstracts by qualifiers, and relations by a number of different features: particles, affixes of case, word order, etc." (Nida, 1975.a:176). There is evidently a clear logic governing such a classification, though it is not difficult to construct an alternative scheme, whilst retaining a similar line of thought with reference to the vocabulary of a language. Nida's universal semantic classification can be adapted for Tamil without much drastic

changes, though one may come across a number of problems while doing so.

4. Information Pertaining to Nouns

Relations pertaining to nouns can be captured by lexical relations such as synonymy, hyponymy, compatibility, incompatibility and meronymy which have been elaborately discussed in the previous section.

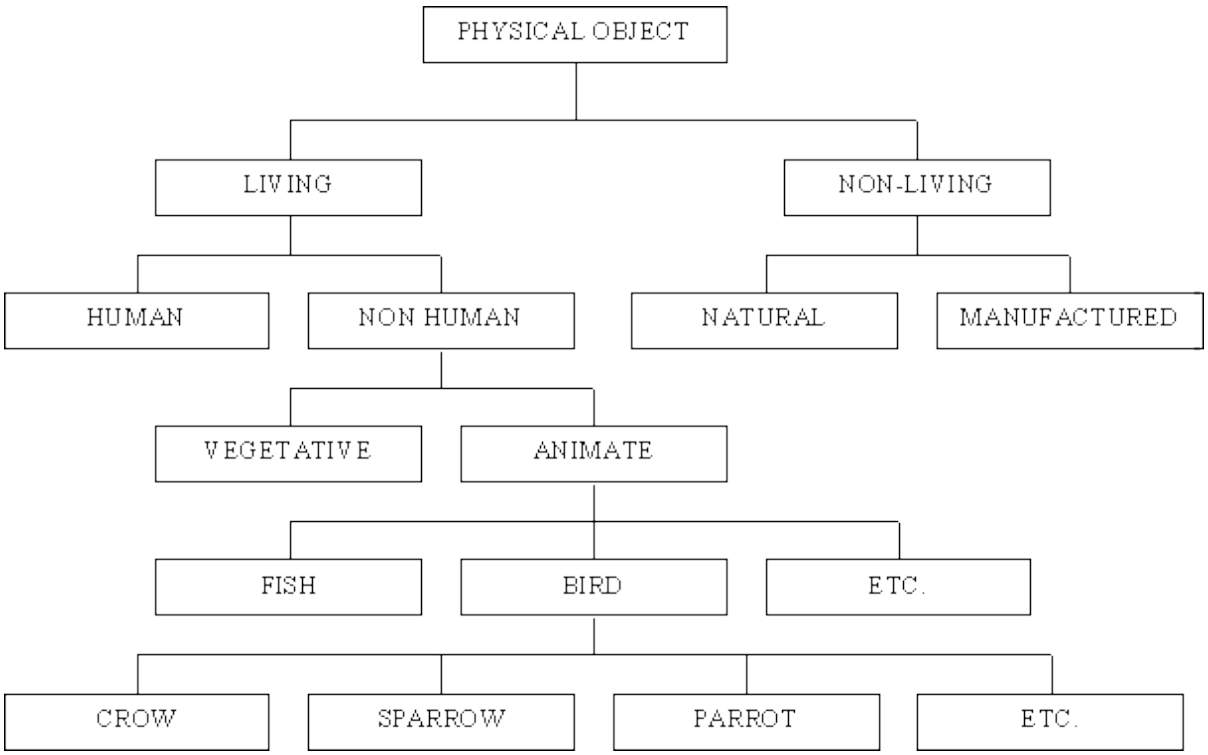
Componential analysis of nouns can help us to group nouns under major semantic domains and subdomains. The dichotomous feature +CONCRETE vs +ABSTRACT helps us to classify nouns into two major groups:

i. Nouns denoting concrete entities
ii. Nouns denoting abstract entities

Another distinguishing feature which helps us to classify noun is the dichotomy between +ANIMATENESS and +INANIMATENESS. Animate nouns can be further classified based on the distinction between flora and fauna. Definitions of common noun typically give a superordinate term plus distinguishing features; that information provides the basis for organizing nouns into domains to create a lexical net work based on their semantic relations. The superordinate relation (hyponymy) generates a hierarchical semantic organization of lexical items. The lexical net is expected to capture the lexical inheritance relation existing between the lexical items as shown in the following example:

uyiruLLavai 'living beings > vilangku 'animal' > paaluuTTi 'mammal' > pacu 'cow'

As far as nouns are concerned it possible to classify them more or less neatly into certain number of semantic domains and subdomains in hierarchical order. The following is a tesauric model for PHYSICAL OBJECT:



5. Information Pertaining to Verbs

The semantic domain EVENTS comprises of verbs. Verbs are arguably the most important lexical and syntactic category of a language. The verb provides the relational and semantic frame-work for its sentence. Its predicate argument structure specifies the possible syntactic structures of the sentences in which it can occur. The linking of noun arguments with thematic roles or cases determines the different meanings of the events or states denoted by the sentence, and the selectional restrictions specify the semantic properties of the noun classes that can flesh out the frame. This syntactic and semantic information generally become part of the verb's lexical entry and thus become part of the information about the verb that is stored in a speaker's mental lexicon. Because of the complexity of this information, verbs are probably the lexical category that is most difficult to study. There are 3312 listed in CreA-vin tarkaallat tamizh akaraati. The list will increase if we take into account the compound verbs.

5.1. Classification Of Verbs

It must be recalled here that Nida's (1976b) tentative classification of events based on componential analyis consists of twelve semantic domains: Physical, Physiological, Sensory, Emotive, Intellection, Communication, Association, Control, Movement, Impact, Transfer, and Complex activities, involving a series of movements or actions. Rajendran (1978) classified verbs into 31 groups out of which ten are major important semantic domains. The important semantic domains identified by him based on componential analysis of verbs are: (1) verbs of movement, (2) verbs of transferring, (3) verbs of change of state, (4) verbs of impact, (5) verbs of senses, (6) verbs of emotion, (7) verbs of intellection, (8) verbs of communication and calling, (9) verbs of association, (10) verbs of cooking. Each major domain is divided into subdomain by taking into account distinguishing semantic component. Say for example, verbs of movement is subclassified into sixteen domains such as verbs of locomotion, verbs of wandering movement, verbs of upward movement, verbs of downward movement, verbs of jumping movement, verbs of circular movement, verbs of movement towards outside, verbs of movement towards inside, verbs of scattering and spreading movement, verbs of shaking movement, verbs of slipping movement, verbs of coming and going, verbs of leaving, verbs of chasing an following, verbs of nearing and approaching, verbs of starting and reaching. Rajendran (1991) classifies verbs into twelve more or less in line with Nida (1976b). The subclassification has been made based on the distinguishing semantic components. The classification may need second look to make it more user friendly. Even though verbs do not show hierarchical ordering, a quasi-hierarchical ordering is possible by taking into account certain pertinent distinguishing features.

5.2. Polysemous Nature of Verbs

The verbs are fewer in number than nouns in Tamil and at the same time verbs are more polysemous in nature than nouns. Verbs can change their meanings depending on the kinds of noun arguments with which they co-occur, where as the meanings of nouns tend to be more stable in the presence of different verbs. Say for example take the verb aTi 'beat', its meaning varies with the nouns to which it is collocated. This kind of semantic flexibility of verbs makes the lexical analysis of verbs difficult.

5.3. Synonymy Among Verbs

Verbal domain exhibit a few truely synonymous verbs. Take for examples the words paTi 'read' and vaaci 'read'. avan puttakam paTikkiRaan 'he is reading a book' can entail avan puttakam vaacoikkiRaan 'he is reading a book'. The relation existing between paTi and vaaci is synonymy and paTi and vaaci are synonyms. at least in this context. Truly synonymous verbs are difficult to find, mostly quasi synonymous verbs are found in Tamil. The existence of a simple and a parallel compound form (noun + verbalizer) causes synonymy (quasi synonymy) in verbal system of Tamil. e.g.

kol 'kill' kolai cey 'murder
vicaari 'enquire' vicaaraNai cey 'investigate'

The synonymous expressions of many verbs show that they are manner elaborations of more basic verbs. For example, viniyooki 'distribute' can be considred as an elaboration of the basic verb koTu 'give'.

5.3. Decompositonal Nature of Verbs

Most approaches to verb semantics have been attempts at decomposition in one form or another. The works of Katz and Fodor (1963), Katz (1972), Gruber (1976), Lakoff (1970) and Jakendoff (1972) and McCawley (1968) stand to testify this point. McCawley (1968), for example, decomposes KILL into CAUSE TO BECOME NOT ALIVE. Jackendoff (1983) have proposed an analysis of verbs in terms of such conceptual categories as EVENT, STATE, ACTION, PATH, MANNER, etc. Relational semantic analysis differs from semantic decomposition primarily by taking lexical items, rather than hypothetically irreducible meaning attoms, as the smallest units of analyis. Thus, relational analysis has the advantage that its units can be thought of as entries in speakers' mental dictionaries. The decompositional features of verbs can be captured partially by the componential features which help in classifying verbs into semantic subdomains.

5.4. Lexical Entailment and Meronymy

The principle of lexical inheritance can be said to underlie the semantic relation between nouns, and bipolar oppositions serve to

organize the adjectives. Similarly different relations that organize the verbs can be cast in terms of one overarching principle, lexical entailment. In logic, entailment, or strict implication, is properly defined for propositions; a proposition P entails a proposition Q if and only if there is no conceivable state of affairs that could make P true and Q false. Entailment is a semantic relation because it involves reference to the states of affairs that P and Q represent. The term will be generalized here to refer to the relation between two verbs V1 and V2 that holds when the sentence Someone V1 logically entails the sentence Someone V2; this use of entailment can be called lexical entailment. Thus for example, kuRaTTai viTu 'snore' lexically entails tuungku 'sleep' because the sentence avan kuRaTTai viTukiRaan 'he is snoring' entails avan tuungkukiRaan 'he is sleeping'; the second sentence necessarily holds if the first one does. Lexical entailment is a unilateral relation: if a verb V1 entails another verb V2, then it cannot be that V2 entails V1. For example uRangku need not entail kuRaTTai viTu.

The entailment relation between verbs resembles meronymy between nouns, but meronymy is better suited to nouns than to verbs. Meronymy (part-whole relation) plays an important role in the hierarchical arrangement of nouns. The division of the human body into parts can serve as a prototype for all part-whole hierarchies.

Fullbaum and Miller (1990) argue that, first, verbs cannot be taken apart in the same way as nouns, because the parts of verbs are not analogous to the parts of nouns. Most nouns and noun parts have distinct, delimited referents. The referents of verbs, on the other hand, do not have the kind of distinct parts that characterize objects, groups, or substances. Componential analyses have shown that verbs cannot be broken into referents denoted solely by verbs. It is true that some activities can be broken down into sequentially ordered subactivities, say for example camai 'cooking' is a complex activity involving a number of sub-activities. Consider the relation between the verbs vangku 'buy' and koTu 'pay'. Although neither activity is a discrete part of the other, the two are connected in that when you buy something, somebody gives it to you. Neither activity can be considered as a subactivity of the other. Consider the relations among the activities denoted by the verbs kuRaTTaiviTu 'snore', kanavukaaN 'dream', and uRanku 'sleep'. Snoring or

dreaming can be part of sleeping, in the sense that the two activities are, at least, partially, temporally co-exensive; the time that you spend snoring or dreaming is a proper part of the time you spend sleeping. And it is true that when you stop sleeping you also necessarily stop snoring or dreaming. The relation between pairs like vaanku 'buy' and koTu 'pay' and kuRaTTaiviTu 'snore' and uRangku 'sleep' are due to the temporal relations between the members of each pair. The activities can be simultaneous (as in the case of vaanku 'buy' and koTu 'pay' or one can include the other (as in the case of kuRaTTaiviTu 'snore' and uRanku 'sleep').

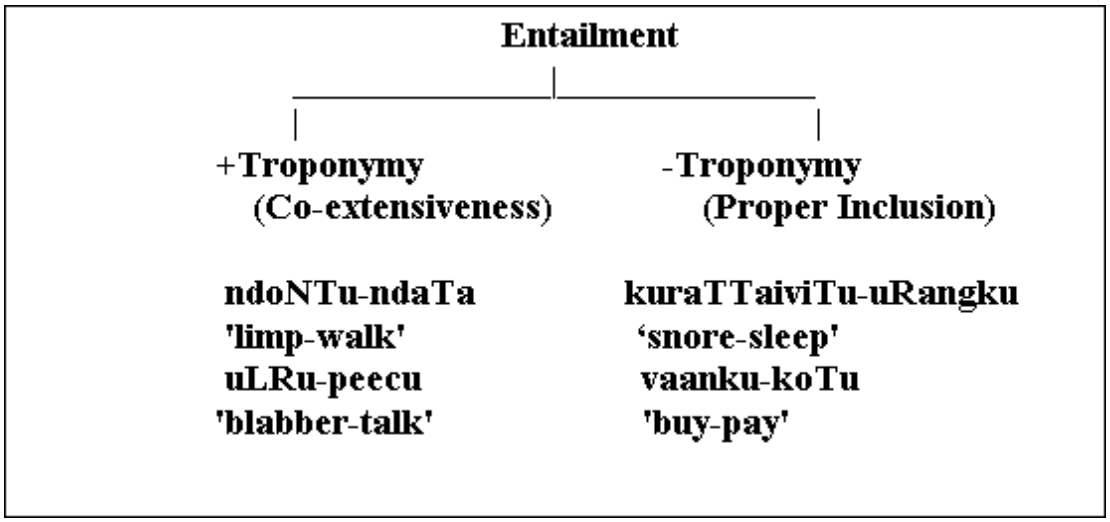
5.5. Hyponymy Among Verbs

Hyponymy is the relationship which exists between specific and general lexical items, such that the former is included in the latter. The set of terms which are hyponyms of same superordinate term are co-hyponyms. The hyponymopus relation of the kind found in nouns cannot be realized in verbs. An examination of 'verb hyponyms' and their superordinates shows that lexicalization involves many kind of semantic elaborations across different semantic fields. The analysis of verbs of motion in Tamil (Rajendran, 1989) reveals the fact that the semantic component such as +DIRECTION (eeRu 'climb up' vs iRangku 'climb down', +MANNER (nazhuvu 'slip down' vs vizhu 'fall' + CAUSE, +SPEED (e.g. uur 'crawl' vs ooTu 'run) added to the common semantic component +MOVE establish co-hyponymous relation found among verbs of motion. Fellbaum and Miller (1990) make use of the term troponymy the establish this type of relation existing between verbs.

5.6. Troponymy and Entailment

Troponymy is a particular kind of entailment in that every troponym V1 of a more general verb V2 also entails V2. Consider for example the pair noNTu 'limp':naTa 'walk'. The verbs in this pair are related by troponymy: noNTu is also naTa in a certain manner. So noNTu is a troponym of naTa. The verbs are also in entailment relation: the statement avan noNTukiRaan 'he is limping' entails avan naTakkiRaan 'he is walking'.

In contrast with pairs like ndoNTu-ndaTa 'limp-walk', a verb like kuRaTTaiviTu 'snore' entails and is included in tuungku 'sleep', but is not a troponym of tuungku. vaangku 'buy'entails koTu 'give', but is not a troponym of koTu 'give'. The verbs in the pairs like kuRaTTaiviTu-tuungku are related only by entailment and proper temporal inclusion. The important generalization here is that verbs related by entailment and proper temporal inclusion cannot be related by troponymy. For two verbs to be related by troponymy, the activities they denote must be temporally co-extensive. The two categories of lexical entailment that have been distinguished so far can be related diagrammatically as shown in the following tree:



5.7. Opposition Relations among Verbs

There is evidence that opposition relations are psychologically salient not only for adjectives, but also for verbs. In building database for verbs, it is found that after synonymy and troponymy, opposition is the most frequently coded semantic relations. The semantics of opposition relations among verbs is complex. As for as Tamil is concerned there is no morphologically derived opposite verbs. Some of the oppositions found among nouns are absent in verbs. A number of binary oppositions have been shown by the verbs, which includes converseness, directional, orthogonal, and antipodal opposition.

5.7.1. Converseness in Verbs

The relation exmplified by the pair husband and wife is known as converseness. Active and passive forms of transitive verbs can be taken as showing converseness opposition. avan avaLai konRaan is in converse relation with the passive expression avaL avanaal kollappaTTaaL. Thus active-passive pairs of transitive verbs in Tamil show converse relation of opposition. By virtue of the definition the converseness, if R is a two-place relation and R' is its converse, we can substitute R' for R and simultaneously transpose the terms in the relation to obtain an equivalence:R(x,y) = R'(y,x). The relation between the verbs vaangku 'buy' and vil 'sell' is rather more complex.

5.7.2. Directional Opposition

The lexical items, which are directionally opposite, are in directional opposition relation. The relationship that hold between the pairs such as vantuceer 'arrive': puRappaTu 'reach', vaa 'come':poo 'go' is directional opposition. Under this category are the verb pairs such as uyar 'rise' and taazh 'go down', eeRu 'ascend' and iRangku 'descend'.

5.7.3. Other Oppositions

There are many other oppositions with reference to change of state, manner, speed, etc. as exemplified below:

kaTTu 'build' : iTi 'demolish'
kaTTu 'tie' : avizh 'untie'
ottukkoL 'agree' : maRu 'disagree'
uLLizhu 'inhale' : veLiviTu 'exhale'
ndaTa 'walk' : run 'run'

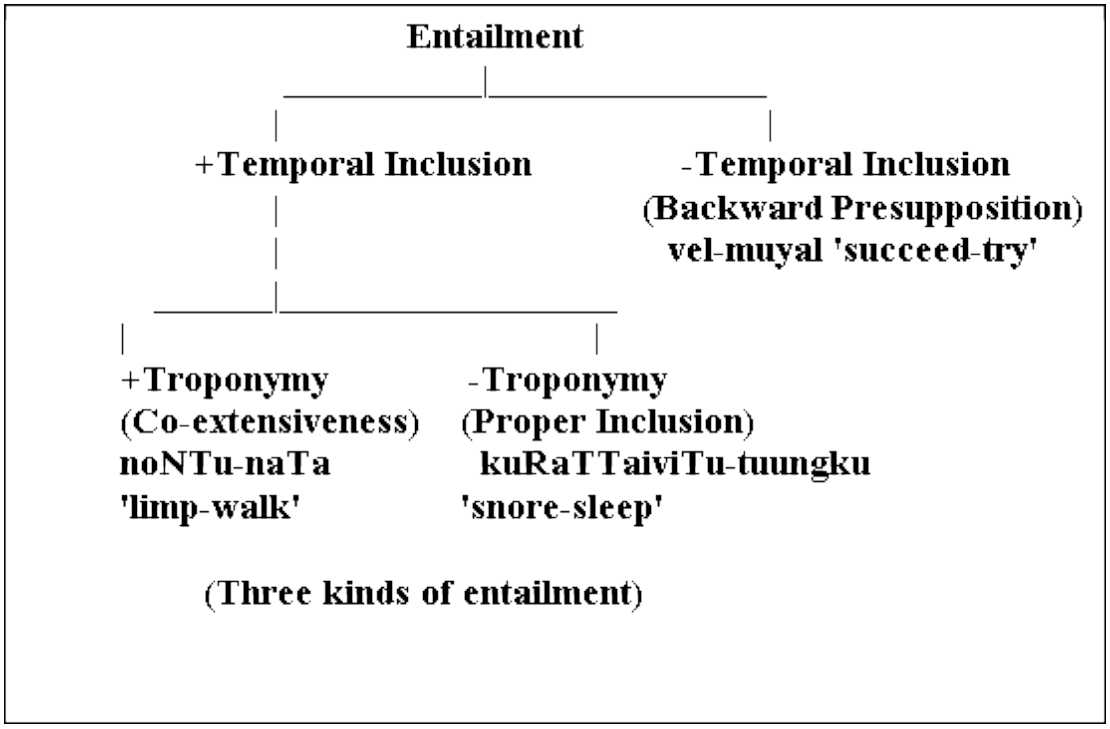
Not only the opposing features, even the presence or absence of a feature can also keep two items in opposition relation. These contrasting or distinguishing features can be arrived at by componential analysis of verbs (Rajendran,1978).

5.8. Opposition and Entailment

Many verb pairs in an opposition relation also share an entailed verb. For example the pair jeyi 'succeed' and tool 'fail' entails muyal 'try'. A verb V1 that is entailed by another verb V2 via backward

presupposition cannot be said to be part of V2. Part-whole statements between verbs are possible only when a temporal inclusion relation holds between these verbs.

The set of verbs related by entailment that we have considered so far can be classified exhaustively into two mutually exclusive categories on the basis of temporal inclusion.



5.9. Causal Relation Among Verbs

The causative relation picks out two verb concepts, one causative (e.g. koTu 'give), the other what might be called the "resultative" (e.g. peRu 'get'). In contrast to the other relations coded in ETTP, the subject of the causative verb usually has a referent that is distinct from the subject of the resultative; the subject of the resultative must be an object of the causative verb, which is therefore necessarily transitive. The causative member of the pair may have its own lexicalization, distinct from the resultative, as in the koTu 'give' and peRu; sometimes, the members of such a pair differ only by a small variation in their common stem, as in the case of vizhu 'fall'-viizhttu 'fell', kaaN 'see'-kaaTTu 'show' . Although many languages have a means to express causation, not all languages lexicalize the causative member independently;

causation is often marked by morpheme reserved for this function. Tamil has a number of lexicalized causative-resultative pairs, such as kaaTTu 'show'- kaaN 'show'; it also has an analytic, or periphrastic causative formed with the auxiliary verbs such as cey and vai added to the infinitive form of the main verb. This is a productive process.

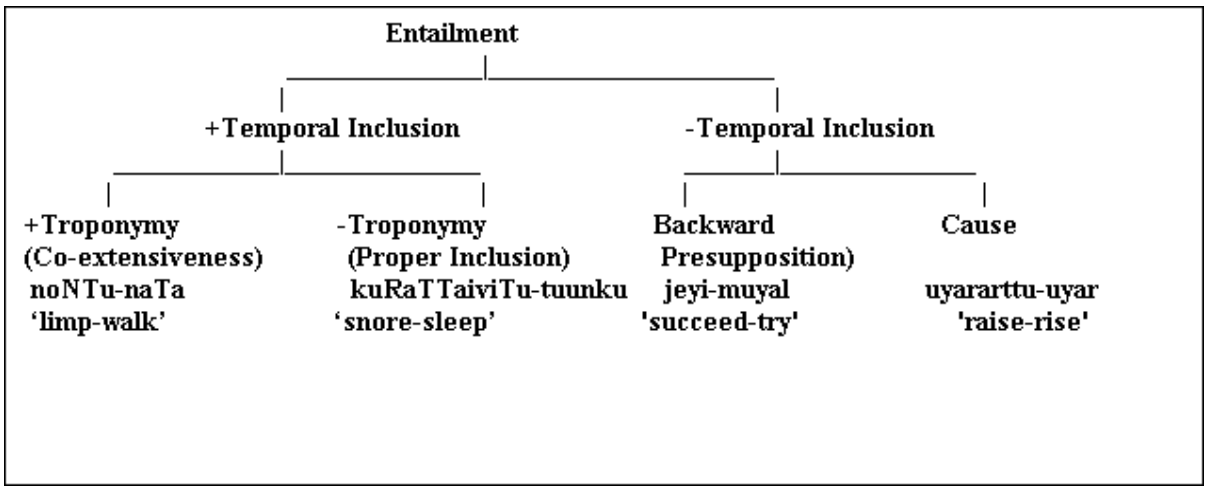
paaTu 'sing' paaTavai, paaTacey 'cause to sing'
tiNaRu 'suffer' tiNaRavai, tiNaRaccey 'cause to suffer'

It has been frequently pointed out that a periphrastic causative is not semantically equivalent to a lexicalized causative, but refers to more indirect kind of causation than the direct lexicalized form. uuTTu 'feed' and uNNaccey 'cause to eat' cannot be interchangeable to refer to the same action, and so are not strictly speaking synonymous expressions of the same concept. For the purpose of ETTP such pragmatic considerations have been disregarded. ETTP recognizes only lexicalized causative-resultative pairs. The synonyms of the members of such pair inherit the cause relation, indicating that this relation holds between the entire concept rather than between individual word form only.

5.12. Causation And Entailment

Causation can be considered as a specific kind of entailment: if V1 necessarily causes V2, then V1 also entails V2. For example the verb veLiyeeRRu 'expel' entails veLiyeeRu 'leave', where the entailing verb denotes the causation of the state or activity referred to by the entailed verb. Like the backward presupposition relation that holds between verbs like tool/vel and muvyal the entailment between verbs like uyarttu 'raise' and uyar 'rise' is characterized by the absence of temporal inclusion.

The causation relation is unidirectional:although veLiyeeRRu entails veLiyeeRu, veLiyeeRu does not entail veLyeeRRu. Similarly uuTTu 'feed' implies uN 'eat', but uN does not entail uuTTu 'feed'.



5.13. Syntactic Properties And Semantic Relations

In recent years there is a trend in incorporating syntactic properties in the lexicon itself. Especially as the verb plays an important role in the interpretation of the sentence in which it forms a part, its syntactic as well as semantic properties are incorporated in the lexicon itself. It is the verb that decides upon the number of noun arguments a sentence should contain. This research analyzes the constraints on verbs' argument-taking properties in terms of their semantic makeup, based on the assumption that a distinctive syntactic behaviour of verbs and verb classes arises from their semantic components.

The ETTP proposed here aims to model lexical memory rather than represent lexical knowledge and so it excludes much of a speaker's knowledge about both semantics and syntactic properties of verbs. There is no evidence that the syntactic behaviour of verbs serves to organize lexical memory. But there is a substantial body of research showing undeniable correlation between a verb's semantic make-up and its syntax, and the possible implications in children's acquisition of lexical knowledge.

To cover at least the most important syntactic aspects of verbs, therefore, ETTP includes for each subdomain of verbs one or several sentence frames, which specify the subcategorization features of the verbs in the synonym set by indicating the kinds of sentences they can occur in. This information permits one quickly to search among the verbs for the kinds of semantic-syntactic regularities that could be established. One can either search for all

the subdomains that share one or more sentences frames in common and compare their semantic properties; or one can start with a number of semantically similar subdomains and see whether they exhibit the same syntactic properties. An exploration of the syntactic properties of co-troponyms occasionally provides the bases for distinguishing semantic subgroup of troponymys.

Viewing verbs in terms of semantic relations can also provide clues to an understanding of the syntactic behaviour of verbs.

6. Thesaurus Information for Adjectives

Though adjectives can be established as a separate grammatical category in Tamil traditional grammarians have taken it partly as verbs and partly as nouns. There are many bound forms or root forms that are purely adjectival in character that cannot be stated as derived from nouns or verbs. The relative participle forms of the verbs, which come before nouns, are also adjectival in their function. So tackling adjectives in ETTP is a challenging problem for Tamil.

6.1.Classification of Adjectives

Dixon (1982) has suggested that the lexical items that are generally found to get included in the category of adjectives can be grouped into seven distinct semantic types. They are:

1. Dimension
kuTTaiyaana 'short'
kuRukalaana 'narrow'

2. Physical Property
periya 'big'
cinna 'small

3. Colour
veLLai 'white'
kaRuppu 'black'

4. Human Propensity
kuruTTu 'blind'
ceviTTu 'deaf'

5. Age
putiya 'new'
pazhaiya 'old'

6. Value
nalla 'good'
keTTa 'bad'

7. Speed
veekamaana 'quick'
metuvaana 'slow'

Adjectives needs to be distinguished into two types: descriptive and relational. Descriptive adjectives ascribe to their head nouns values of bipolar attributes and consequently are organized in terms of binary oppositions (antonymy) and similarity of meaning (synonymy). Descriptive adjectives that do not have direct antonyms are said to have indirect antonyms by virtue of their semantic similarity to adjectives that do have direct antonyms. Cross references has to be maintained between descriptive adjectives expressing a value of an attribute and the noun by which that attribute is lexicalized. Reference-modifying adjectives have special syntactic properties that distinguish them from other descriptive adjectives. Relational adjectives are assumed to be stylistic variants of modifying nouns and so are cross-referenced to the noun files.

All languages provide some means of modifying or elaborating the meanings of nouns, although they differ in the syntactic form that such modification can assume. Tamil syntax allows a variety of ways to express the qualification of a noun. For example, if naaRkaali 'chair' alone is not adequate to select the particular chair a speaker has in mind, a more specific designation can be produced with adjectives like periya 'large' and vacatiyaana 'comfortable'. Words belonging to other syntactic categories such as relative participle form of verbs and nouns can function as adjectives.

Past participle form as adjectives
iruNTa viiTu 'dark house'
mankiya oLi 'dim light'
varaNTa nilam 'dry land'

Nouns as adjectives
talaimai atikaari 'chief officer'
tiruTTu paNam 'illegal money'
iNai aayvaaLar 'co-inverstigator'

Nouns phrases as well as clauses can modify a noun.

avanuTaiya taattaavin ndaaRkaali 'his grandfather's chair'
neRRu kaTaiyil vaankiya ndaaRkaali 'the chair which was bought from the shop yesterday'

Nouns modification is primarily associated with the syntactic category 'adjective'. Adjectives have as their sole function the modification of nouns, whereas, modification is not the primary function of noun or prepositional/postpositional phrases. The lexical organization of adjectives is unique to them, and differ from that of the other major syntactic categories, noun and verb.

The adjective domain in ETTP contain mostly adjectives, although some nouns and relative participial forms of verbs that function frequently as modifiers have to be entered as well.

6.2. Descriptive Adjectives

Descriptive adjectives are what one usually thinks of when adjectives are mentioned. A descriptive adjective is one that ascribes a value of an attribute to a noun.

atu kanamaana cumai 'that luggage is heavy' presupposes that there is attribute eTai 'WEIGHT' such that eTai + cumai = kanam 'heavy'. Similarly low and high are values of HEIGHT. A thesaurus should have cross reference between descriptive adjectives and the domain of noun that refer to the appropriate attributes.

The semantic organization of descriptive adjectives is entirely different from that of nouns. Nothing like the hyponymic relation that generates nominal hierarchies is available for adjectives: it is not clear what it would mean to say that one adjective 'is a kind of' some other adjective. The semantic organization of adjective is more naturally thought as an abstract hyperspace of N dimensions rather than as a hierarchical tree.

6.2.1. Antonymy in Descriptive Adjectives

The basic semantic relation among descriptive adjective is atonymy. The importance of antonymy first became obvious from results obtained with word association tests. The importance of antonymy in the organization of descriptive adjective is understandable when it is recognized that the function of these adjectives is to express values of attributes, and that nearly all attributes are bipolar. Antonymous adjectives express opposing values of an attribute. For example, the antonym of kanamaana 'heavy' is ileecaana 'light' that expresses a value at the opposite pole of the WEIGHT attribute. This binary opposition is to be represented in ETTP.

6.2.2. Gradation in Descritpive Adjectives

Most discussions of antonymmy distinguish between contradictory and contrary terms. This terminology originated in logic, where two propositions are said to be contradictory if the truth of one implies the falsity of the other and are said to be contrary if only one proposition can be true bout both can be false. For example uyiruLLa 'living' and uyiraRRa 'non-living' are contradictory terms as atu uryiruLLa jantu 'it is a living creature' necessarily implies atu uyiraRRa jantu alla 'it is not a non-living creature'. But kuNTaana 'fat' and melinta 'thin' are contrary terms because maalaa kuNTaana peN and maalaa melinta peN connot both be true, although both can be false if malaa 'Mala' is of average weight. Contraries are gradable adjectives, whereas contradictions are not. Gradation must also be considered as semantic relation organizing lexical memory for adjectives. The following data will exemplify the gradation found among nouns:

kotikkiRa 'very hot'
cuuTaana ' hot'
iLanjcuuTaana 'warm'
kuLirnta 'cold'

ETTP will account for the gradation found among adjectives.

6.2.3. Markedness in Adjectives

Most of the attributes have an orientation. It is natural to think of them as dimensions in a hyperspace, where one end of each dimension is anchored at the point of origin of the space. The point of origin is the expected or default value; deviation from it merits comment, and is called the marked value of the attribute.

For example, the antonyms akalamaana 'wide' and kuRukkalaana 'narrow' can illustrate this general linguistic phenomenon known as markedness.

atu nuuRu miiTTar akalamaana caalai 'the road is 100 meters width'
atu akalamaana caalai 'that is a wide road'
*atu pattumiiTTar kuRukalaana caalai 'the road is 10 meters narrow'

Thus the primary member, akalamaana 'wide' is unmarked term; the secondary member, kuRukalaana is marked.

6.2.4. Polysemy and Selectional Preferences

Polysemy is found among verbs as a limited number of adjectives are used to attribute a considerable number of nouns. For example, the use of ndalla in the following phrases will illustrate the polysemous nature of it.

ndalla kaalam 'good time'
ndalla naaNayam 'good coin'
ndalla ndaNpan 'good friend'
ndalla ceuppu 'good chappal'

The semantic contribution of adjectives is secondary to, and depends on, the head noun that they modify.

Adjectives are selective about the nouns they modify. The general rule is that if the referent denoted a noun does not have attribute whose value is expressed by the adjective, then the adjective-noun combination requires a figurative or idiomatic interpretation. For example, a road can be long because roads have LENGTH as an attribute, but stories do not have LENGTH, so ndiiNTa 'long' does not admit literal readings. The semantic contribution of adjectives

is secondary to, and dependent on, the head nouns that they modify.

6.2.5. Syntax of Adjectives

In Tamil, the premodifying form of the adjective is different from the postmodifying form, as the premodifying form is a real adjective whereas the postmodifying form is a pronominalized noun. Take for example for the adjective form ndalla 'good' which comes before a noun and the predicative forms are nallavan 'good male person', ndallavaL 'good female person' ndallavar 'good person', ndallatu 'good thing' depending on the noun which is attributed.

avaL ndalla peN 'she is a good woman'
peN ndallavaL 'the woman is good'

ETTP should account for the relation between the pure adjectives and their derived nominal forms.

The predicative form can be of the following structure too:

N+ aaka + iru

aaka is an adverbial equivalent of the adjectival formative suffix aana. iru is a be-verb. The structure helps to relate a simple adjectival form to its predicate form. The following example will illustrate the point.

avaL azhakaana cirRumi
she beauty_become_RP girl
'she is a beautiful girl'
ciRumi azhakaanavaL 'the girl is beautiful'
ciRumi azhakaaka irukkiRaaL 'the girl is beautiful

6.3. Reference-Modifying Adjectives

Distinction can be drawn between reference modifying and referent-modifying adjectives. For example pazhaiya 'old' in the phrase en pzhaiya ndaNpan'my old friend' does not refer the refetent who is a person as old, but attributes the friendship as old. Where as pazhaiya in pazhaiya paattiram 'old vessel', the adjective

pazhaiya attributes directly the vessel itself. Similarly in the phrase,

ndeeRRaiya kuRRavaaLikaL inRaiya mandtirikaL
yesterday criminals today's ministers
'Yesterday's criminals are today's ministers'

both the adjectives attributes the quality of being criminals and the quality of being ministers respectively, rather than the persons. Some reference modifying adjectives resemble descriptive adjectives in that they have direct antonyms: ndeeRRaiya 'past'/innaaLaiya 'present', muntaiya'past'/inRaiya 'present'

6.4. Colour adjectives

Colour terms are organized differently than other adjectives. They can be both nominal as well as adjectival. They are adjectives as they can be graded and conjoined with other descriptive adjectives. But the pattern of direct and indirect anotonymy that is observed for other descriptive adjectives does not hold good for colour adjectives.

6.5. Relational Adjectives

Another kind of adjective comprises the large and open class of relational adjectives. Relational adjectives mean something like 'of, relating/pertaining to or associated with' some noun, and they play a role similar to that of a modifying noun. For example, cakootara 'fraternal', as in cakootra paacam 'fraternal love' relates to cakootran 'brother', and poruLaataara 'economical', as in poruLaataara eeRRa taazhvu 'economical difference', is related to poruLaataaram 'economics', As for as Tamil is concerned noun form is used mostly in the place of relational adjective in English. For example,

icaikkruvi 'musical instrument' paRcuttam 'dental hygiene'

Since relational adjectives do not have antonyms, they cannot be incorporated into the clusters that characterize descriptive adjectives. And because their syntactic and semantic properties are a mixture of those of adjectives and those of nouns used as noun modifiers, rather than attempting to integrate them into either

structure ETTP maintains a separate file of relational adjectives with cross references to the corresponding nouns.

7. Designing And Implementation Of ETTP

7.1. Problem Of Organization

The fact that the lexical items of a language fall into a number of distinct parts of speech would itself preclude the ordering of the vocabulary of a language in terms of hyponymy and hypernymy. But it is difficult to assign members of a part of speech under a superordinate term or even a set of items under a superordinate term. "If we include quasi-hyponymy with hyponymy as a relation in terms of which vocabularies are structured hierarchically, the hypothesis that the vocabulary in all languages is structured hierarchically under a relatively small set of lexemes of very general sense is rather possible. It is this hypothesis, however, which is difficult to evaluate on the basis of the evidence that is at present available." (Lyons: 1977:299)

The lexical relations synonymy, hyponymy, compatibility, incompatibility and meronymy are widely talk about in the context of thesaurus and dictionary and it can be taken for granted that the lexicographers are aware of the problems involved in deciding the items that can be grouped by these relations.

7.2. Superordinate Term

Finding suitable superordinate term or heading for a group of items often poses problem. The availability of a superordinate term, which can cover a semantic domain, facilitates the assigning of membership to the domain. Often there may be no superordinate form to include a group of items. The non-availability of superordinate term can be managed by a number of techniques: by promoting a specific term into a super ordinate term, by using descriptive phrases, etc.

7.3. Establishing Semantic Domains And Sub Domains

The hierarchical structuring of lexical items and establishing boundary between domains can sometimes be difficult. One cannot always expect a sharply defined and dichotomously distinguished

neat classification. Also, though it may be easy to decide about the inclusion of items that form the core of a domain, there may be many borderline cases where the decision of inclusion becomes difficult.

7.4. Bringing Out The Net Work Of Lexical Relations

An electronic thesaurus should bring out the semantic relations existing between the lexical items. Though it may be difficult to show all these relations in a paper thesaurus, it is possible to bring out these relations in an electronic thesaurus for the benefit of its users in a friendly fashion.

7.5. Arrangement Of Lexical Items

Though arrangement of lexical items under a terminal domain is not a serious issue in ETTP, the decision of what to give first among a set of synonyms is a problem to be encountered in the preparation of the ETTP. Criteria based on core meaning vs. peripheral meaning, higher style vs. lower style and more frequently used vs. less frequently used may help in deciding the priority in the arrangement of synonyms.

7.6. Relating Basic Meaning With Derived Meaning

Derivational meanings can be established through basic meaning, i.e. derived lexical items can be related to the basic items form which they are derived. This helps in assigning componential features to the derivatives from the componential features of bases, thus avoiding duplication. Duplication can be avoided by cross-reference. Listing the derived items under the class to which they belong by virtue of their derived meanings and furnishing with cross-reference for the bases from which they are derived seem to be a better solution. This can be easily accomplished in an ETTP.

7.7. Computational Issues

A database has to be created for the development of ETTP. If dictionaries and thesaurus are available in the electronic media, they can be used for the purpose of developing a database for ETTP.

7.7.1. Computer Corpus

A computer corpus, as we understand, is a large body of naturally occurring computer-readable texts or text extracts used for research, and especially for the development of natural language processing. Any work related to lexicon presumes a computer corpus these days. Using a computer corpus for lexical work makes things easy. Also the corpus helps in making any decision with maximum perfectness. A corpus annotated for grammatical categories and semantic information will be a useful tool in thesaurus making. Automatic tagging for grammatical information is feasible, but automatic tagging for semantic information needs a knowledge base, the building up of which is a difficult task. Building a separate corpus for the preparation of thesaurus is a welcome thing, though one can make use of an already available Tamil corpus, for example that which is prepared by the CIIL, Mysore under their project **Development of Corpora of Texts of Indian Languages in Machine Readable form**.

7.7.2. Creation Of A Database

Corpus can be made use of for the creation of database for making a ETTP. A tagged corpus can help us to prepare a thesaurus more effectively. Dictionaries can be used as a secondary source. The paper thesaurus prepared by Rajendran (2001) can be used to create the data base for the preparation of ETTP. The hierarchical classification of vocabulary of Tamil is available in the above mentioned paper thesaurus.

7.7.3. Semi-Automatic Way Of Classifying The Lexical Items

What is needed form a corpus for the thesauric classification of lexical item into semantic domains is semantic information in terms of the semantic relations such as synonymy, hyponymy, compatibility and incompatibility. The lexical relations like meronymy and metonymy have also to be taken into consideration for a wider perspective.

TEXTS ----- > PROCESSING ----- > SEMANTIC INFORMATION

Apart from the conceptual information or information concerning the possible relations between meanings of lexical items, the following information are also welcome:

1. Definitional information or information concerning the internal meaning structure of lexical items
2. Contextual information or information concerning the meaning of a lexical item with reference to the context in which it is used
3. Collocation information or information concerning the meaning of a lexical items in sequence

Automatic tagging, annotating and coding the text for the above mentioned semantic information is a difficult task. This can be partially achieved through an efficient semantic parser that is constructed to bring out the network of meaning relations existing between lexical items. A morphological analyzer and syntactic parser can help to categorize the lexical items in terms of their grammatical functions in sentences. But semantic information requires intelligence on the part of computer, which can be achieved partially incorporating artificial intelligence.

7.7.4. Expanse of ETTP

One's first impression of ETTP is likely to be that it is an on-line thesaurus. It is true that sets of synonyms are basic building blocks, and with nothing more than these synonym sets the system would have all the power of a thesaurus. When short glosses are added to the synonym sets, it resemble as on-line dictionary that has been supplemented with synonym sets, it resembles an on-line dictionary that has been supplemented with synonymous for cross referencing. But ETTP includes much more information than that. It will have synonym sets, hypernym-hyponym sets, meronym-holonym sets, troponym sets, entailment sets, scale-sets, serial sets, cyclic sets, rank-sets, and other relevant multi-member sets. The relations between hypernym and hyponym, meronym and holonym, and binary oppositions such as dirctional opposition, antipodal opposition, orthogonal opposition, converseness, compementarity, etc. will be captured by efficient programs. In an attempt to model the lexical knowledge of a native speaker of Tamil, ETTP has been given detailed information about relation between word forms and

synonym sets. How this relational structure should be presented to a user raises questions that outrun the experience of conventional lexicography.

In developing this on line-lexical database, it has been convenient to divide the work into two interdependent tasks, which bear a vague similarity to the traditional tasks of writing and printing a dictionary. One task was to write the source files that contain the basic lexical data; the contents of those files are the lexical substance of ETTP. The second task was to create a set of computer programs that would accept the source files and do all the work leading ultimately to the generation of a display for the user. The ETTP system falls naturally into four parts: the ETTP lexicogrpahers' source files, the software that converts these files into the ETTP lexical database, the ETTP lexical database, and the set of software tools used to access the database.

7.7.5. Lexical Source Files

ETTP source files will contain all sorts of lexical information. They will be output of a detailed componential and relational analysis of lexical items. A variety of lexical and semantic relations are used to represent the organization of lexical items. Two kinds of building blocks need to be distinguished in the source files: word forms and word meanings. Word forms will be represented in their familiar orthography; and their meaning relations will be represented in terms of lexical sets like synonym sets, hyponym sets, meronym sets, cyclic sets, serial sets, scale sets, rank sets, etc. Two kinds of relations will be recognized: lexical and semantic. Lexical relations hold between word forms; semantic relations hold between word meanings. ETTP will organize nouns, verbs, adjectives and adverbs into lexical sets, which include the synonym sets, which will be further, arranged into a set of lexicographer's source files by syntactic category and other organizational criteria. Nouns, verbs, adjectives and adverbs are grouped according to semantic fields.

Each source file will contain a list of lexical sets for one part of speech. Each lexical set consists of semantically related word forms, relational cross references, and other pertinent information. The relations represented by these cross-references include (but are limited to): hypernymy/hyponymy, antonymy, entailment, and

meronymy/holonymy. Polysemous words are those that appear in more than one synonym sets, therefore representing more than one concept. The Compiler utility will complies the lexicographers' files.

7.7.6. Cross Reference

Cross-reference relates a word form in a set and to a related form in another sets. Lexical relations exist between relational adjectives and the nouns that they relate to, and between adverbs and the adjectives form which they are derived. The semantic relation between adjectives and the nouns for which they express values are encoded as attributes. The semantic relation between noun attributes and the adjectives expressing their values are also encoded. Antonyms are lexically related. Synonymy of word forms is implicit by inclusion in the same synonym set. Meronymy can be further specified as a part of something or a substance of something or a member of some group. Holonymy can also be specified in the same manner, each cross-reference representing the semantic relation opposite to the corresponding meronymy relation. Many cross-references are reflexive, that is, if a set contains a cross reference to another set, the other set should contain a corresponding reflexive cross-reference back to the original set. The Compiler can automatically generate the relations for missing reflexive cross-references.

7.7.7. Storage System

The lexical source files will be maintained in a Storage System based on a Control System for managing multiple revisions of text files.

7.7.8. Compiler System

The function of the Compiler System is to compile primarily the lexicographer's files into a database format that facilitates machine retrieval of the information in ETTP. The Compiler has several options that control its operation on a set of input files. To build a complete ETTP, all of the lexicographers' files must be processed at the same time. The Compiler is also used as a verification tool to ensure the syntactic integrity of the lexicographers' files when they are returned to the storage system with the restored command.

7.7.9. Morphological Analyser and Generator

Many dictionaries hand their information on uninfected headwords without separate listings for inflectional (or may derivational) forms of the word. In a printed dictionary, that practice cause little trouble; with a few highly irregular exceptions, morphologically related words are generally similar enough in spelling to the reference form that the eye, aided by boldface type, quickly picks them up. In an electronic dictionary, on the other hand, when an inflected form is requested, the response is likely to be frustrating announcement that the worked is not in the database; users are required to know the reference form of every word they want to look up. In ETTP, only base forms of words are generally represented. In order to spare users the trouble of affix stripping a Morphological Analyser (MA) is introduced in the whole set up so that the dictionary form can be arrived at automatically. For replacement we may require an inflected equivalent of the text form and this can be achieved by making use of Morphological Generator (MG) in the set up.

MA and MG handle a wide range of morphological transformations. MA uses two types of processes to try to convert a word form into a form that is found in the ETTP database. There are list of inflectional endings, based on syntactic category that can be detached from individual words in an attempt to find a form of the word that is in ETTP. There are also exception lists for each syntactic category in which a search for an inflected form may be done. MA tries to use these two processes in an intelligent manner to translate the word form passed to the form found in ETTP. MA first checks for exception, then uses the rules of detachment. In the case synonym replacement, MG will take up the selected synonym and generate its inflected form equivalent to the text.

7.7.10. Retrieval of Lexical Information

In order to give a user access to information in the database, an interface is required. Interfaces enable end users to retrieve the lexical data and display it via window-based tool or the command line. When considering the role of interface, it is important to recognize the difference between a printed dictionary and a lexical database. ETTP's interface software creates its responses to a

user's requests on the fly. Unlike an on-line version of a printed dictionary, where information is stored in a fixed format and displayed on demand, ETTP's information is stored in a format that would meaningless to an ordinary reader. The interface provides a user with a variety of ways to retrieve and display lexical information. Different interfaces can be created to serve the purpose of different users, but all of them will draw on the same underlying database, and may use the same software functions that interface to the database files.

7.7.11. User Friendly Interface

A user-friendly interface helps the user to get the needed information with great ease. This includes the following issues:

a. Transfer of classificatory and network scheme into commands for electronic type setting,
b. Transfer of classificatory and network scheme into information retrieval system, and
c. On-line retrieval: combining the classificatory and network scheme with code searching.

REFERENCES

Atkins, B.T.S. and A. Zampolli. 1994. ***Computational Approaches to the Lexicon***. Oxford: Oxford University Press.

Beckwith, R. Miller, G.A. 1993. ***Design and Implementation of the WordNet Lexical Database and Searching Software***.

Bolinger, D. 1967. Adjectives in English: attribution and predication. ***Lingua***, 18.1-34.

Cruse, D.A. 1986. ***Lexical Semantics***. Cambridge: Cambridge University Press.

Dixon, R.M.W. 1982. ***Where have all the adjective gone?*** Berlin: Mouton Publishers.

Graside, Roger, Geoffrey Leech & Geoffrey Samson (eds.) 1987. ***The Computational Analysis of English: A Corpus based Approach***. London: Longman.

Gruber, J. 1976. ***Lexical Structures in Syntax and Semantics***. New York: North Holland.

Guckler, G. 1983. Appendix: B: A Computer-based Monolingual Dictionary: A Case Study. In R.R.K. Hartmann (ed.) 1983. ***Lexicography: Principles and Practice***. London: Academic Press Inc.

Jackendoff. 1972. ***Semantic Interpretation in Generative Grammar***. Cambridge, Mass.: MIT press.

Jones, K.S. 1986. ***Synonymy and Semantic Classification***. Edinburgh: Edinburgh University Press.

Katz, J.J. 1972. ***Semantic Theory***. New York: Harper and Row.

Katz, J.J. and Fodor, J. 1963. 'The Structure of Semantic Theory'. ***Language*** 39:170-210.

Levin, J.N. 1989. ***Towards a Lexical Organization of English verb***. Ms., Evanston: Northwestern University.

Lyons,J.1977. ***Semantics***, 2 vols. New York: Cambridge University Press.

Martin, W.J.R., B.P.F.Al and P.J.G.Van Sterkenburg. 1983. On Processing of A text Corpus. In R.R.K. Hartmann (ed.). 1983.***Lexicography: Principles and Practice***. London: Academic Press Inc.

McCawley, J.D. 1968. 'Lexical Insertion in a Transormational Grammar without Deep Structure. Darden, B.J., Bailey C-J.N, Davison (eds.)1968. ***Papers from the Fourth Regional Meeting***. Chicago Ill.: Department of Linguistics, University of Chicago Pp 71-80.

Miller, G.A. 1993. ***Nouns in WordNet: A Lexical Inheritence System***.

Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D, and Miller, K. 1993. ***Introduction to WordNet: An On-line Lexical Database***.

Fullbaum, C., Gross, D. and Miller, K. 1993. ***Adjectives in WordNet***.

Nida, E.A. 1975a. ***Compositional Analysis of Meaning: An Introduction to Semantic Structure***. The Hague: Mouton 1975.

----------b. ***Exploring Semantic Structure***. The Hague: Mouton.

Pandey. M.K. 1995. An Electronic Thesaurus: Theoretical Premise. In Francis Ekka et al (ed.) 1995. ***Indian Congress of Knowledge and Language*** vol. I. Mysore: CIIL.

Rajendran, S. 1978. ***Syntax and Semantics of Tamil Verbs***. Ph.D. Thesis. Poona: University of Poona.

---------- 1981. "Semantic Structure of Tamil verbs". In: 13th All India Tamil Teachers' Association Conference ***aayvukkoovai*** vol. 2.305-310.

---------- 1882. "Verbs of Seeing in Tamil". In: ***Bulletin of the Deccan College Research Institute***, Pune.

---------- 1983. "Coming and Going in Tamil". In: ***To Greater Heights***. Mysore: CIIL.

-------- 1995. "Towards a Compilation of a Thesaurus for Modern Tamil". ***South Asian Language Review***. 5.1:62-99.

--------- 1996. "The Feasibility of Preparing a Thesaurus using Corpus". Workshop on Indian Language Corpus and its Applications (28, 29 Oct. 1996). Central Institute of Indian Languages, Mysore.

--------- 1997. "Intricacies Involved in the Verbnet for Tamil". DLA Conference Telugu University, Hyderabad.

---------- 2001. ***taRkaalattamizh coRkaLanciyam [Thesaurus for Modern Tamil]***. Thanjavur: Tamil University.

---------- 2002. **"Preliminaries to the preparation of wordnet for Tamil".** ***Language in India***, 2:1, www.languageinindia.com

---------- 2002. **Semantic structure of Directional verbs of movement in Tamil.** ***Language in India***, 2:6, www.languageinindia.com

--------- 2002. **Syntax and semantics of Verbs of transfer in Tamil**. ***Language in India***, 2:8, www.languageinindia.com